TRANSIT EQUITY IN PITTSBURGH:

A data-driven approach

2023

TABLE OF CONTENTS

OUR TEAM

We are a group of master's students at Heinz College, Carnegie Mellon University, with diverse academic backgrounds:

Mahnoor Ayub
Cailin De Zeeuw
Yuxuan He

Matt Lampl
Sara Usha Maillacheruvu
Sai Tej Reddy

ACKNOWLEDGEMENTS

We extend the warmest gratitude to everyone who made this project possible:

Sara Khalil
David White
Dr. Lee Branstetter
Cameron Drayton
Dr. Xiaodong Qian
Bike Share Pittsburgh

POGOH

Carnegie Mellon University
Heinz College
INFORMATION SYSTEMS • PUBLIC POLICY • MANAGEMENT

EXECUTIVE SUMMARY

Bike Share Pittsburgh operates POGOH, Pittsburgh's bikeshare network. In the process of expanding its network, Bike Share Pittsburgh works hard to carefully site new stations to improve transit equity in the city. By enhancing network equity, the organization aims to make it easier for low-income Pittsburghers to access jobs, grocery stores, and more.

We developed a data-driven tool to support Bike Share Pittsburgh in its decision-making process around siting new stations. Building off research by Dr. Xiaodong Qian, a transportation expert, we implemented a three-part model that optimizes new station locations to maximize accessibility. Here, we define accessibility as the difference in the number of grocery stores, jobs, and other opportunities that a rider can access by moving from one station to the other. Alongside its existing community engagement efforts, Bike Share Pittsburgh can use this tool to help ensure that its network expansion facilitates greater transit equity in the region.

Image Source: POGOH

In this report, we provide a conceptual overview of our modeling process, as well as the key findings that emerged from our analysis and proposed station locations from our models. In the appendix, readers can find technical details on our modeling process.

Lastly, we emphasize that our data-driven approach be taken as an entry point into Bike Share Pittsburgh's siting process. In conjunction with input from riders and the broader community, Bike Share Pittsburgh can build a network that increases transit equity in Pittsburgh.

GLOSSARY

Throughout this report, we use language that may be unfamiliar to some readers. Here, we provide an overview of key terms that readers will encounter.

Disadvantaged area	Designated as such if classified as "High Need" or "Very High Need" per Allegheny County's Community Need Index
Trip origins	How many trips **start** at a given station.
Trip destinations	How many trips **end** at a given station.
OD pair	Origin-Destination pair for a given trip. E.g., 21-33 would indicate the trip started at station 21 and ended at station 33.
Opportunities	Number of hospitals, grocery stores, jobs, etc. someone can access in the Census block group where a station is located.
Accessibility Change	Difference (positive, negative) in opportunities generated by moving from one station to another. **Our proxy for transit equity**. E.g., moving from a disadvantaged area to an advantaged area would result in an increased accessibility score.
K value	The number of new stations to be chosen by the genetic algorithm.(E.g., k = 5 would select 5 optimal stations out of the 238 candidates.)Specified by the user based on how many output stations are desired.
MJM	The Mobility Justice Membership is a $10 annual bikeshare pass for eligible Pittsburgh residents. It is designed "to make bikeshare accessible to persons whose income limitations prevent them from using bikeshare at its standard pricing."

INDUSTRY OVERVIEW

Roam Bikeshare

Omaha, Lincoln, and Valentine, NE

Recently rebranded, ROAM Share is a nonprofit bikeshare organization that is the product of recent mergers across three Nebraska cities: Heartland Bike Share in Omaha, BikeLNK in Lincoln, and Valentine Bike Share in Valentine. Resulting in a network of nearly 300 bikes, 54 kiosks, and 16 staff members, the merger has enabled ROAM to work toward economies of scale in leveraging its physical infrastructure as well as human capital. Aiming for fully electricized operations, ROAM emphasizes the need for fleet modernization in growing its ridership: the organization introduced e-bikes in 2018 and noted a 150% increase in usage.

ROAM's history is marked by a series of rebrands and mergers. In 2011, a nonprofit in Omaha began operating a small bikeshare program through Bcycle, a bikeshare company owned by Trek Bicycle that currently operates in over 30 US cities, including Indego in Philadelphia and Houston BCycle. While BCycle is a private company, it works with local governments and organizations to implement bikeshares. The operation became Heartland Bike Share in 2013 and expanded in 2017 when it was contracted to provide bikeshare services by two additional Nebraska clients. They later expanded to Lincoln, NE — home of the flagship campus of the University of Nebraska network — after a feasibility study and grant greenlit the expansion.

CoGo Bikeshare

Columbus, OH

CoGo Bike Share System was launched in Columbus, Ohio, in July 2013 with 300 bicycles and 30 stations located throughout downtown Columbus, but has since expanded to include 90 stations and 600 bikes serving Columbus, Bexley, Upper Arlington, Grandview Heights, and Easton. E-bikes were introduced in June 2020, making it more accessible and convenient for users. CoGo's station network provides twice as many docking points as bicycles to ensure that there is always an available return dock. The City of Columbus owns the CoGo bikes and stations and made the initial investment in the system, but the network is currently operated by Lyft. The Central Ohio Transit Authority (COTA) and the Mid-Ohio Regional Planning Commission (MORPC) have also realized the importance of the bike-sharing program, and the City has been in discussions with those entities to join in sharing program costs.

In 2022, there were nearly 54,000 trips taken on a CoGo bike, a record for the system. CoGo is continually seeking sponsors to help expand and improve its reach, and business owners are encouraged to provide employees with memberships or discounts to use the system or sponsor a station and display their company logo as a grassroots supporter.

INDUSTRY OVERVIEW

Madison BCycle
Madison, WI

Madison, Wisconsin's bikeshare program, Madison BCycle, is known for being the first fully electrified city bike fleet in the United States. With more than 300 electric-assist bicycles (or "e-bikes") and 50 stations, Madison's bike share program offers insights into the implications of expanding e-bike access. BCycle's e-bike pilot program resulted in increased trip counts, compared to standard "pedal-powered" bikes. The availability of e-bikes has reportedly prompted riders to take longer trips. This, in turn, has sparked a regional expansion of BCycle docking stations in the metropolitan area.

Cost per ride spans from $5 for a single ride pass, to $25 for a monthly pass, to $150 for an annual pass. The monthly and annual passes allow unlimited 90-minute rides, with an additional $3 charge for each additional 30 minutes. Discounts to membership are offered to students at the nearby University of Wisconsin - Madison, as well as Madison College. Additionally, a partnership with the Madison Public Library Foundation provides BCycle passes available for check-out at the local library. Each of the nine library locations offers two Community Passes, which can be checked out for up to a week by library patrons. As of 2023, it appears that Madison no longer offers a mobility justice membership option.

MetroBike Austin
Austin, TX

MetroBike Austin comprises 75 BCycle stations, offering a mix of electric and pedal-powered bicycles. The MetroBike system is owned by the City of Austin in partnership with its public transit system, CapMetro, and the bike share program is operated by the non-profit Bike Share of Austin. The system has served the City of Austin since 2013, but underwent a major expansion in 2022. MetroBike Austin boasts of being "one of the most successful bike sharing programs in the country" due to record-breaking ridership during the South by Southwest music festival.

MetroBike passes can be purchased through the CapMetro App or the BCycle App. Standard ride length is 60 minutes, before additional charges are incurred. Pass costs range from the unlocking fee of $1.09 for the Pay-as-you-ride option, to the annual membership at $86.60 for unlimited rides all year. Discounts are offered to students at some local universities, as well as to low-income residents that meet specific eligibility requirements. BCycle for All offers $5 annual passes for those who qualify.

INTRODUCTION

About Bike Share Pittsburgh

Image Source: POGOH

Bike Share Pittsburgh is the nonprofit that owns and operates the bikeshare system in Pittsburgh known as POGOH. Their mission is to "provide Pittsburgh with a joyful, sustainable, and affordable mobility service for all residents and visitors."

POGOH launched in May 2022 with 37 stations, equipped with a fleet of 350+ bikes including traditional pedal bikes and e-assist bikes. Around 20 of the docking stations are also charging stations, connected to street light electricity through a unique partnership with Duquesne Light. Demonstrating their focus on equity, POGOH debuted a Mobility Justice Membership – an annual membership pass for just $10, available to people who participate in government assistance programs. POGOH bikes offer Pittsburgh residents a micromobility option, which fits into the larger landscape of public transit in Pittsburgh. For a more detailed exploration of Pittsburgh Public Transit, see Appendix 1.

Project Motivation

Bikeshare systems can provide the benefits of bicycling without the burdens of ownership. They can benefit the physical health and mobility of individuals, as well as add social and environmental benefits at the community level. Shared bikes also represent a significant opportunity to replace short private vehicle trips (0–3 miles) and reduce transportation sector emissions. According to the US Department of Transportation, the average U.S. household produces about 9.5 trips a day, about half of which are within three miles. E-bikes have the potential of replacing almost every 1 in 5 of these short trips. Bike-sharing can also reduce congestion to upwards of 4% within a neighborhood.

Despite the benefits, increased popularity, and expansion of various bikeshare systems, disadvantaged populations are not highly representative of current bikeshare users' profiles. There are multiple types of barriers, such as financial barriers (e.g., lack of credit card), cultural barriers, and safety concerns, among others. Bike Share Pittsburgh is dedicated to working with BIPOC communities and low-income neighborhoods to better connect residents to the bikeshare system and improve transit equity in Pittsburgh.

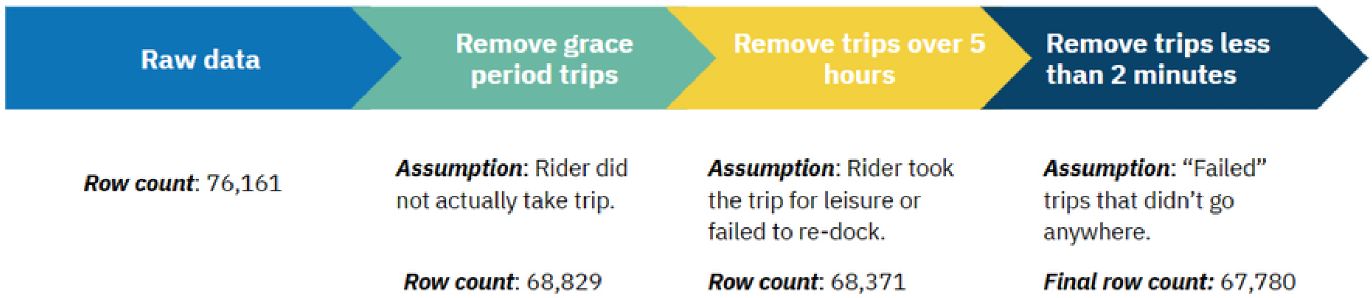
Project Objectives

Using a data-driven approach, we have developed a tool, along with a network plan, that uses transportation equity and mobility justice as guiding influences to site new bikeshare stations. The model estimates the potential demand (i.e., bikeshare trip origins and destinations) and its distribution. It then evaluates performance over a set of objectives (e.g., accessibility improvements) to find the optimal distribution of stations that can:

- Improve service in communities with pre-existing stations through increased network density.
- Expand the network into new neighborhoods.
- Increase ridership, particularly among the disadvantaged communities.

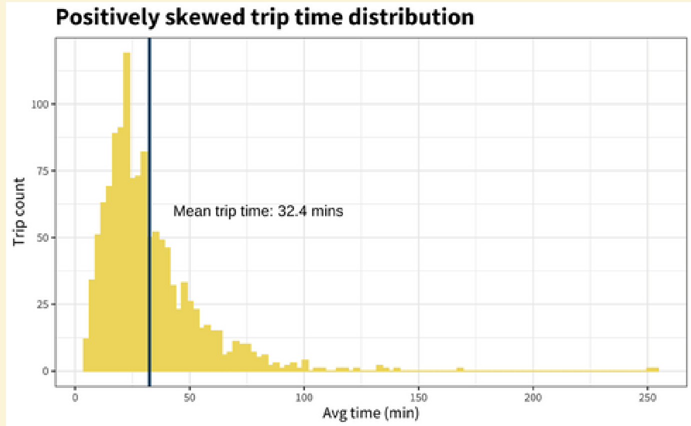
DATA EXPLORATION

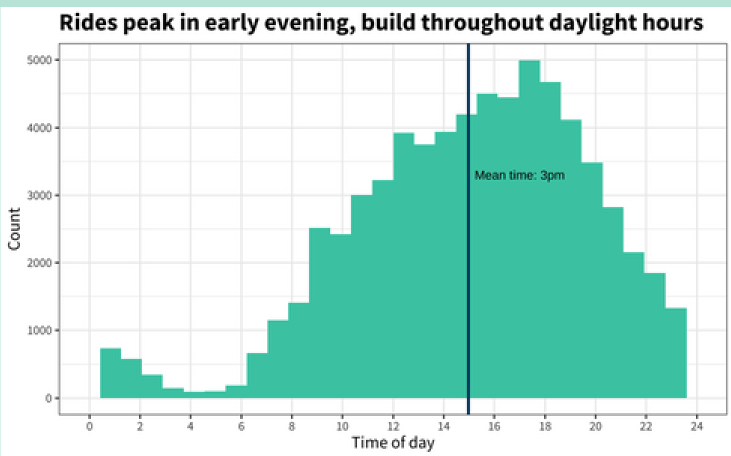
Before diving into the modeling process, we explored the data across 3 key dimensions: Time, Geography, and Ridership. Across these analyses, we focused on riders participating in the Mobility Justice Member (MJM) program, which offers a discounted membership ($10/annually) to participants in social service programs, such as SNAP. MJM riders offer a crucial lens into how "disadvantaged" riders use and move throughout POGOH's network. This is especially important given one of the key assumptions of our modeling process: because we do not have person-level demographic information, we assume that riders who originate a trip in a given Census block group have the "average" characteristics (e.g., income) of that block group.

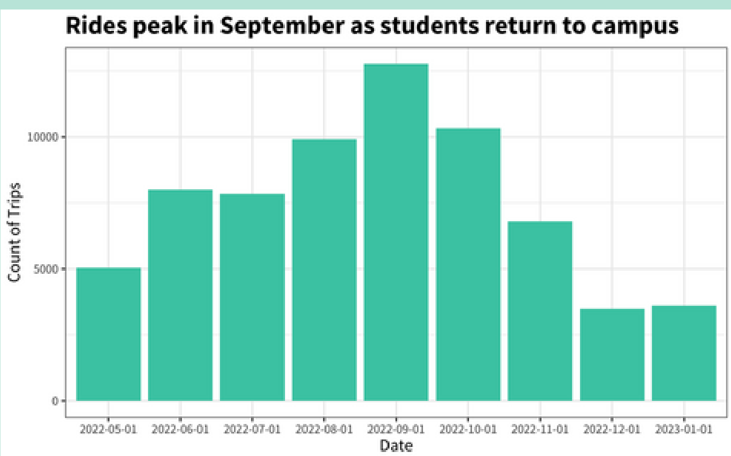


DATA EXPLORATION: TIME

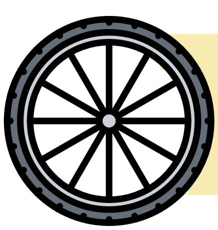
Here, we see that most POGOH riders take short trips, with 75% of trips ending under 30 minutes. The mean trip time is 32.4 minutes, which falls close to the free-first-30-minutes cutoff that POGOH members receive.







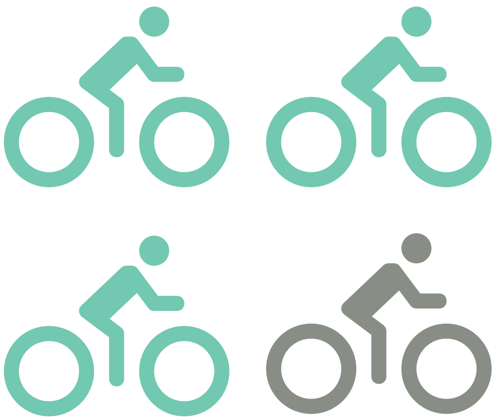
We also analyzed the time of the day in which POGOH riders prefer to ride. We see that daily ridership peaks around 6pm, with rides building throughout the daily hours and falling as the sun sets. Additionally, the data suggest that the most popular time of year is in August, September, and October, which coincides with the start of the school year for university students. With the onset of the winter months, ridership begins to decline.



Takeaway: Most riders take short trips. Ridership peaks in the fall. The early evening is the most popular time to ride.

DATA EXPLORATION: RIDERSHIP

Annual members take 3 out of 4 trips



Next, we turn to ridership data. There are 11,419 unique riders in the dataset, and most riders take few trips: the mean number of trips taken is 5.94 trips in the 9 months of data available to us (May 2022 to January 2023). In fact, nearly 40% of all riders take only 1 trip.

We also see that POGOH members drive the overall trip count: nearly three in four trips are taken by POGOH members, with the remaining quarter taken by non-members. Across the sample period, the highest trip count per rider was 717 trips, or roughly 80 trips per month.

While MJM riders make up a small fraction of the overall ridership — comprising less than 1.1% of POGOH's ridership — they take a disproportionate number of trips. In fact, they take nearly five times more trips than their ridership share. On average, MJM riders take 32.4 trips, while non-MJM POGOH members take 10.3 trips per rider, across the sample period. MJM riders leverage POGOH three times more than non-MJM members, suggesting that bikeshare can be an important mode of transportation for some disadvantaged populations.

Rider type	Ride count	Percent of riders
Non-member	18,523	27.3%
Member	49,257	72.7%

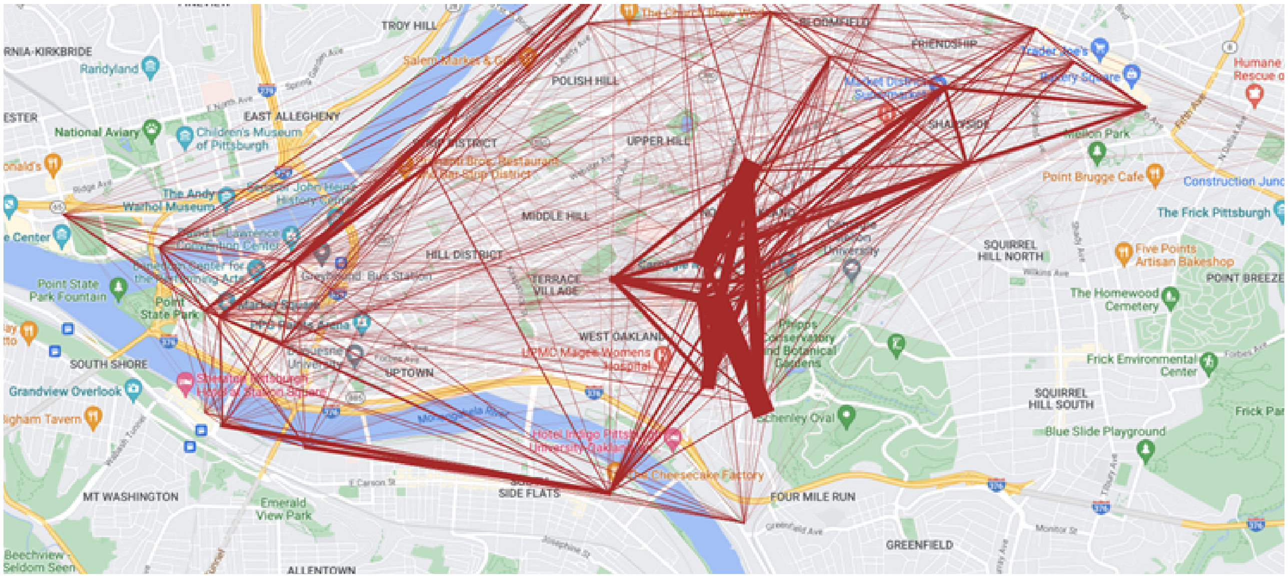
Rider type	Ride count	Percent of riders
MJM riders	3,338	4.9%
Non-MJM	64,442	95.1%

Rider type	Rider count	Percent of riders
MJM riders	97	1.1%
Non-MJM	11,322	98.9%

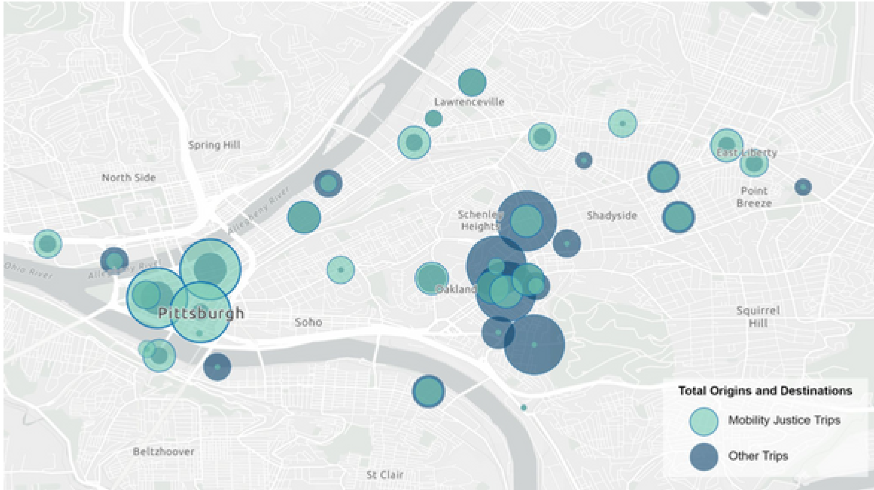
Takeaway: MJM riders take more trips than non-MJM riders. Members drive overall trip counts.

DATA EXPLORATION: GEOGRAPHY

The diagram below traces trips across POGOH's existing network. We can see that trips tend to concentrate in Oakland, as well as downtown, with other high-frequency nodes around the Shadyside and Southside Works areas.



There are important differences to highlight between the preferred routes of MJM riders and non-MJM riders. MJM riders prefer routes centered in downtown, as well as the Hill District. Non-MJM riders, on the other hand, prefer Oakland-centric routes.



These differences are likely driven by the fact that University of Pittsburgh students receive a free membership, which concentrates non-MJM routes around the campus.

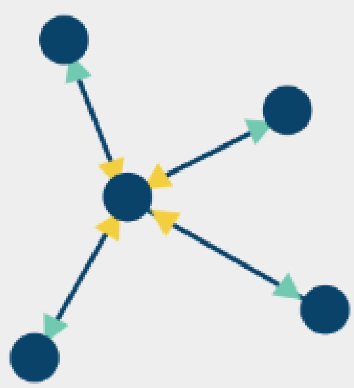
Takeaway: The vast majority of trips are centered around Oakland. MJM riders tend to start and end their trips in downtown, while other riders generally ride in and around Oakland.

OUR APPROACH

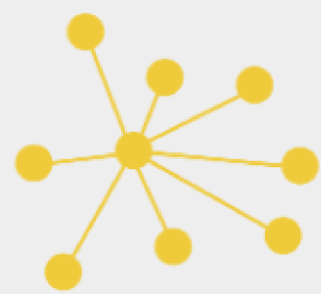
Drawing from Dr. Xiaoding Qian's methodology, we follow a three-phased modeling process, with the ultimate goal of understanding where to site new stations to maximize overall network accessibility, measured as the change in the number of opportunities (grocery stores, hospitals, etc.) that someone can access by moving between two stations.

Trip Origins & Destinations

Network Flow



Optimizing Station Network



Phase I is Trip Origins and Destinations, where the goal is to predict the number trips that will start at a given station, as well as the number of trips that will end at a given station.

Phase II is Network Flow, where the trip distribution across the entire network is predicted. This phase attempts to predict the number of trips taken from station A to station B, for all origin and destination pairs in the network.

Phase III is Optimizing Station Network. Given a list of candidate station locations, this phase chooses K number of stations that will maximize accessibility in the entire network.

PHASE I:
TRIP ORIGINS AND DESTINATIONS

In Phase I, we estimate ridership in a block group, given its neighborhood characteristics. We create two models — one for origins and one for destinations — using the negative binomial regression model. This model is similar to the more commonly known linear regression model but is good for outputs that involve count data. Because we predict the number of trips that start and end at each station, this model works well for the task at hand.

To estimate ridership for origins and destinations for each station in the network, we collected data at the block group level. Thus, our predictions are also at the block group level (i.e., the number of predicted trip origins/destinations will be the same for all stations in a block group). This means that the geographic granularity of our analysis will be at the level of the block group; we are not predicting, for example, the exact cross street where a station should be located in Phase III.

Our approach uses eight key predictor variables, shown in the table to the right, that were identified by Qian and Jaller (2020) as significant predictors of ridership. Again, we aggregated these variables at the block group level before running them through the models. The resulting output is the number of trip origins and the number of trip destinations in a given block group.

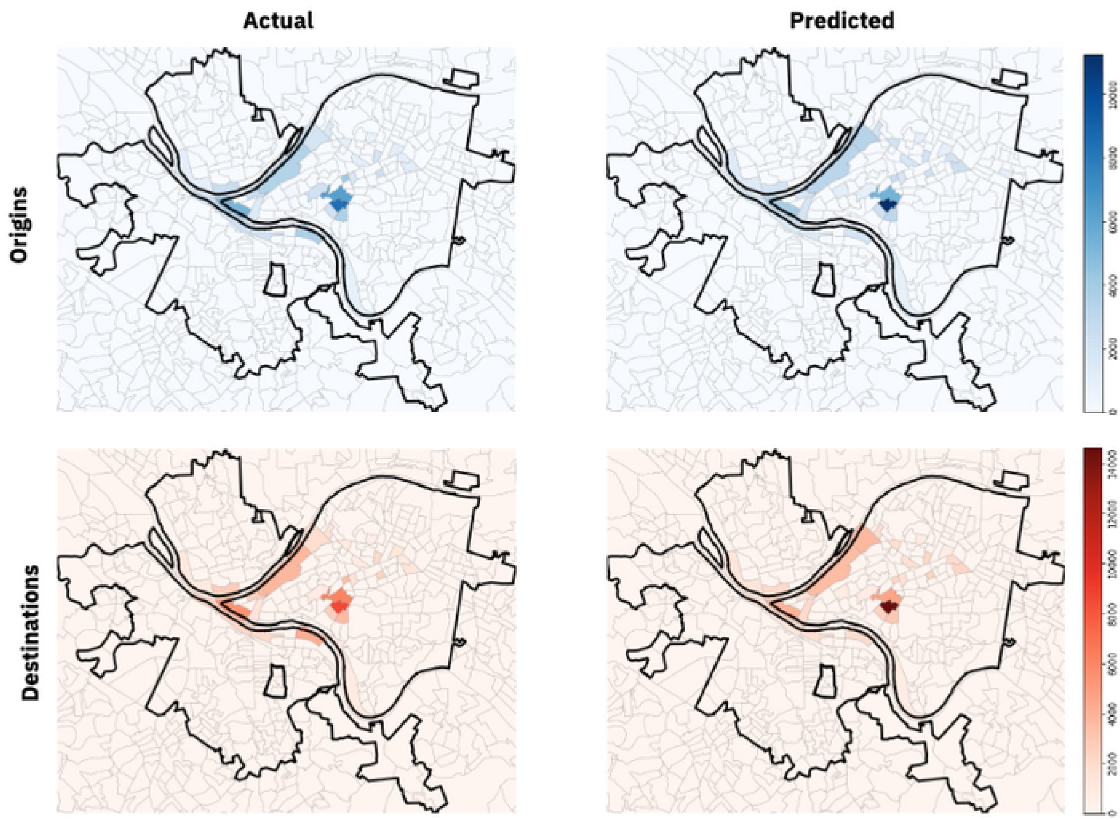
Variable	Description
Labor Force	Number of people who work in a given block group
Percent of Young Population	Population between the ages of 20 and 35
Employment Rate	Percent of working-age population that is employed
Bike Path Density	Percent of total area of block group covered by bike infrastructure (bike lanes, paths, etc.)
Park Areas	Total area of block group covered by a park
Number of Transit Stops	Count of PRT transit stops in a block group
Stations within block group	Number of existing POGOH stations within block group
Disadvantaged indicator	Binary indicator if a block group falls within an area identified as High or Very High Need by the Community Needs Index

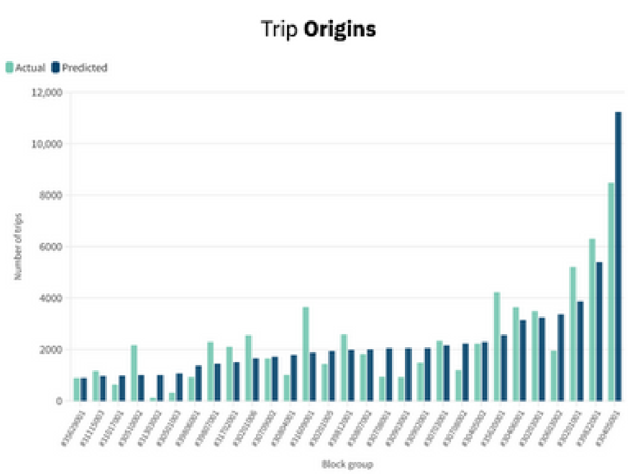
Takeaway:

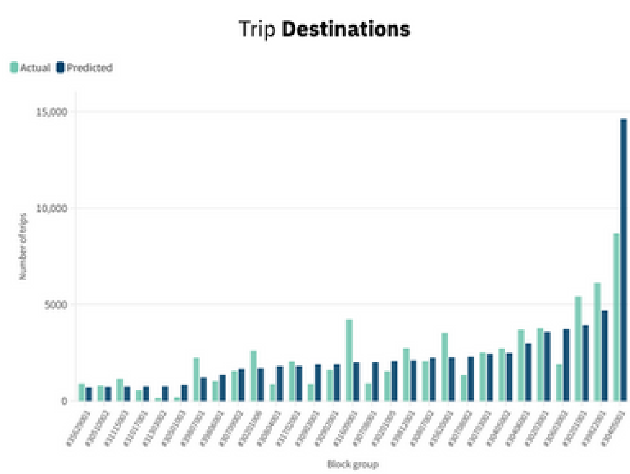
In Phase I, we use demographic data and past POGOH trip data to predict the number of trips that will start and end at a given station. This is an input to our next phase.

PHASE I: RESULTS
TRIP ORIGINS AND DESTINATIONS

As the maps and graphs below demonstrate, we yield results with strong predictive power with our Phase I model, well-approximating trip origins and destinations of POGOH's actual network. This sets up nicely for Phase II, in which we determine the network flow. For additional discussion of this model, including more technical nuance, please refer to Appendix 2.







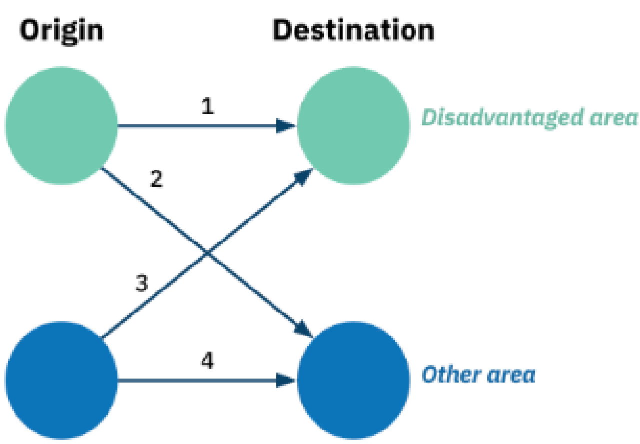
PHASE II:
NETWORK FLOW

In Phase II, we estimate the distribution of trips in the network — in essence, how trips flow through the network. In Phase I, we predicted the number of trips that start and end at each station in the network. In Phase II, we use those origin and destination predictions to estimate how many trips will be taken between any two given stations. This helps us to understand destination choices and behaviors of bikeshare users.

Following a competing destination model based on a gravity model, we aim to estimate the number of trips originating in a given station to all other stations in the network. We also assess the effect of improved accessibility on destination choices for different user types — members and non-members. The model assumes that there will be more trips between stations with higher gains in accessibility (i.e., more opportunities like grocery stores, schools, jobs, etc.) and less trips between stations that are located far from one another.

To calculate the gains in accessibility, we used data from the Longitudinal Employer-Household Dynamics dataset as well as points of interest (groceries, schools, and hospitals) available in a given station's block group. Next, to calculate the travel time between each pair of stations, we used the Google Maps API. Trip time estimates from Google Maps also captures the elevation gain or loss between two stations. For example, if station A is lower in elevation than station B, then the travel time will be longer from A to B, than it would be from B to A.

Using these inputs, along with the observed number of trips between two given stations, we estimated two parameters: rho, a measure of increased accessibility, and beta, the willingness to travel between two areas. These two parameters were estimated for both member riders and non-member rides, as well as for each OD-pair type. A positive rho indicates an increase in accessibility, while a positive beta indicates that there is a higher willingness to travel between two areas. The findings for these parameters are shown in the table on the following page.



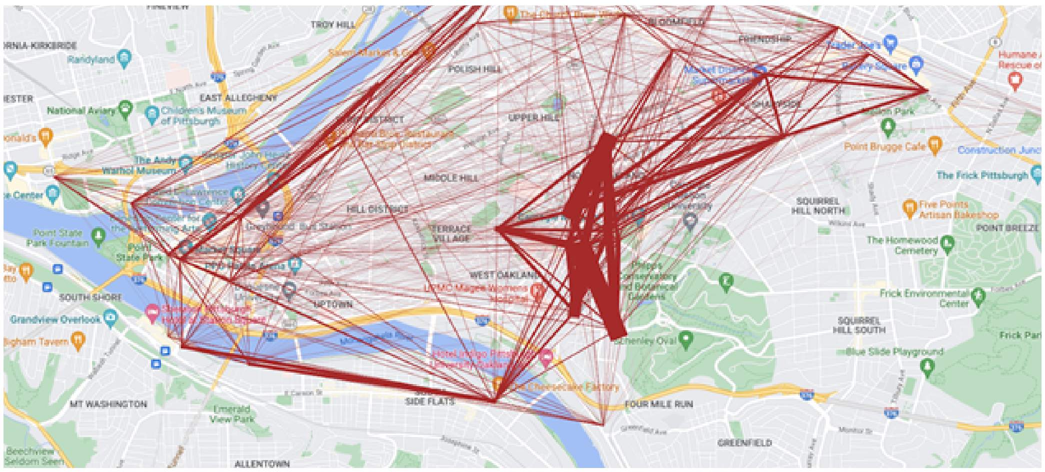
Takeaway:

In Phase II, we take the predicted number of trips that start and end at each station and determine how they *flow* across the network, to other stations within the network. This helps us to understand the behaviors of bikeshare users.

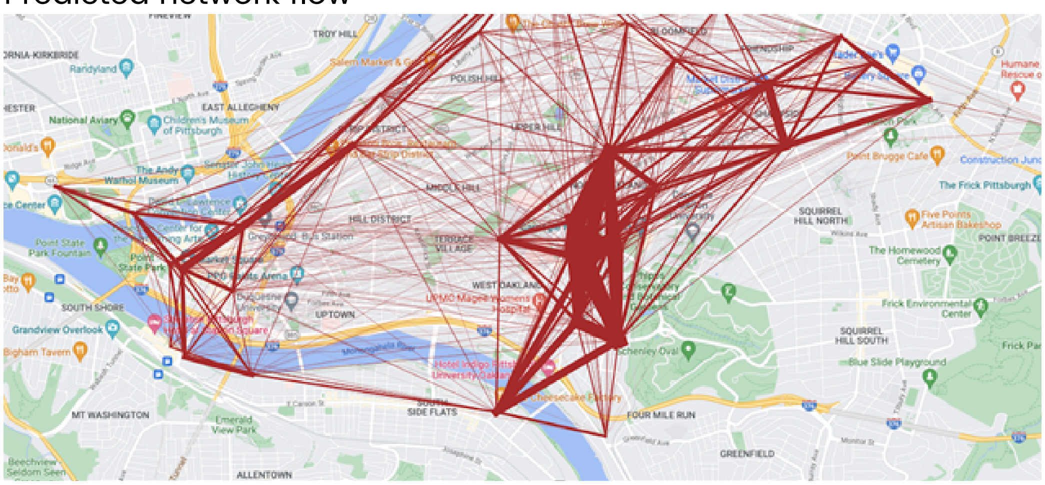
PHASE II: RESULTS
NETWORK FLOW

With these parameters set, we are then able to estimate the flow for the entire network. The actual trip distribution, as well as the predicted trip distribution, are shown below. Here, we can see that the predicted network flow well approximates the actual network flow. This will serve as an input to our final model, where we optimize the location of new stations given rider behaviors and network patterns.

Actual network flow



Predicted network flow



OD Pair Type	Description	Rho (non-member)	Beta (non-member)	Rho (member)	Beta (member)
1	Disadvantaged-Disadvantaged	-0.02	-0.76	0.11	-0.48
2	Disadvantaged-Other	0.01	-0.64	0.09	-0.52
3	Other-Disadvantaged	0.02	-0.65	0.10	-0.49
4	Other-Other	0.02	-0.68	0.09	-0.50

PHASE III:
OPTIMIZING STATION NETWORK

The next step in our modeling process focuses on siting new stations across Pittsburgh to maximize accessibility, defined as the change in the number of opportunities (jobs, grocery stores, etc.) that someone can access by moving from one station to another. It's important to note that we are optimizing for accessibility and not, for example, for the increase in trips originating in disadvantaged areas and ending in advantaged areas. In turn, optimal site stations may not actually be located in disadvantaged areas.

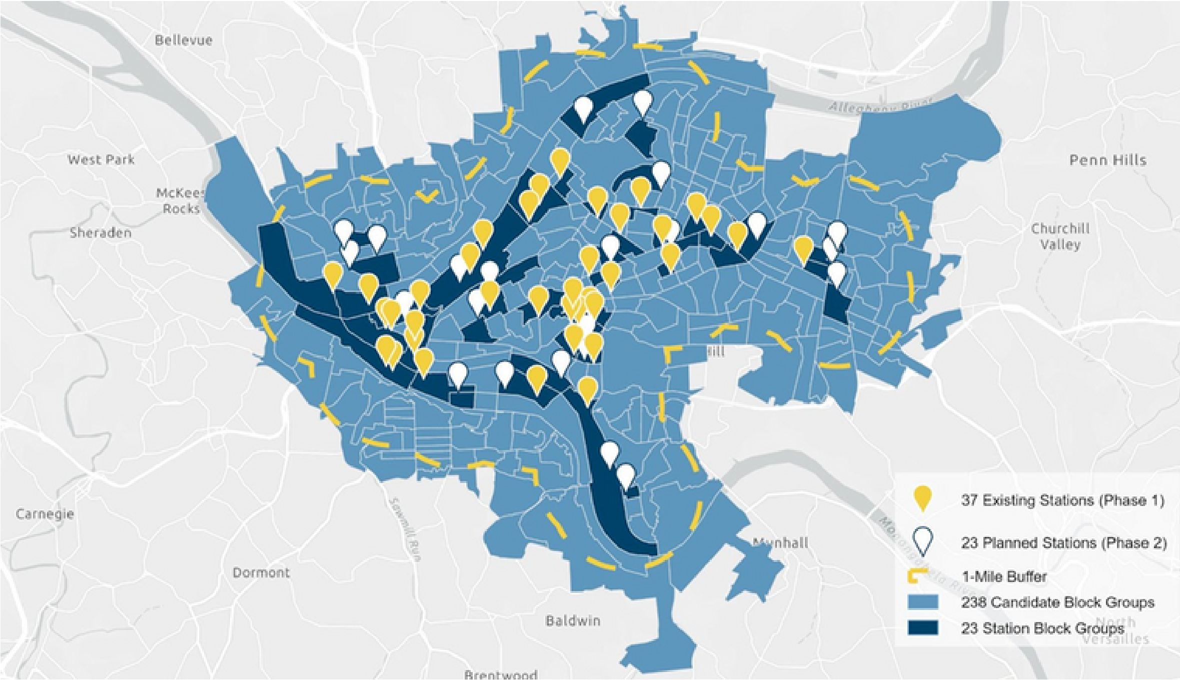
Following Qian, et al (2022), we leveraged the genetic algorithm to determine optimal station locations. The genetic algorithm is an evolutionary algorithm that optimizes the total network at one time, accounting for synergies between stations. It is "genetic" in nature because it uses evolutionary techniques (selection, crossover, and mutation) to produce an optimal or near-optimal solution. For more discussion of the technical workings of this model, please reference the Appendix 4.

Inputs to the algorithm include the network flow results generated from Phase II — how trips move throughout the network — as well as a list of candidate block groups where new stations could potentially be sited. The candidate block groups include all groups within a 1-mile radius of current POGOH stations. While we could have considered all Allegheny County block groups, best station siting practice is iterative, with new stations sited relatively close to existing stations to generate more trips. In addition, by narrowing the block group list geographically, we reduced computational cost, enabling the algorithm to converge more quickly.

After running the algorithm, we generated lists of optimal locations to site new stations, maps of which are included in the Recommendations section that follows. These locations maximized overall network accessibility, resulting in the greatest increases in opportunities for riders.

Takeaway:

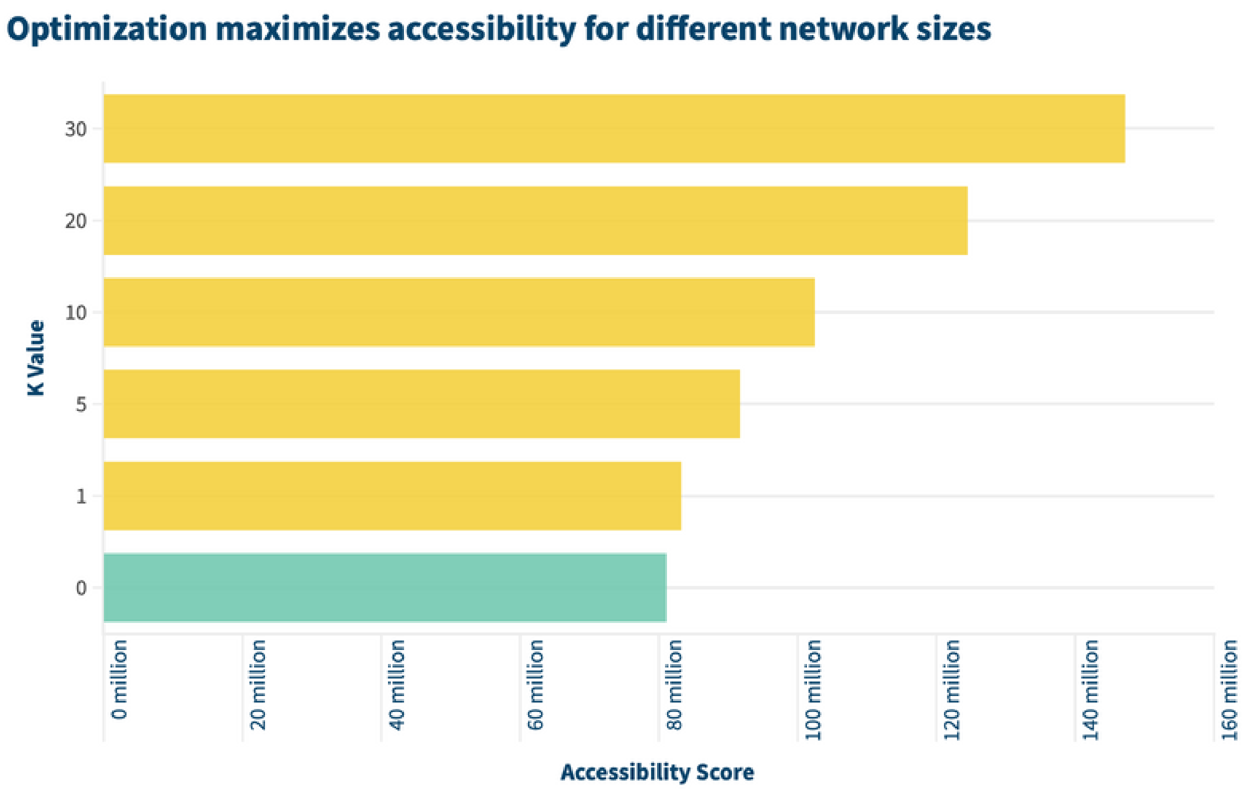
In Phase III, we optimize for *accessibility,* the change in opportunities between two stations, using the genetic algorithm.



RECOMMENDATIONS

Based on our exploration and analysis of the data, we propose several recommendations for Bike Share Pittsburgh. Our primary recommendations are in the form of optimal station locations sited to maximize equity. In addition, we have several broader suggestions for Bike Share Pittsburgh to consider.

As previously discussed, the optimization model maximizes accessibility for a given network size. For our recommendations, we chose to produce five different optimal network configurations. First, we recommend the optimal station location if POGOH were to add just one station to the existing network. Next, we recommend the optimal station locations for adding five stations at once. Then, for 10 stations, 20 stations, and 30 stations.



Summarized above are the five different runs of the model, each using a different K value between 1 and 30. We compare the accessibility score of each proposed solution to the score of the existing network.

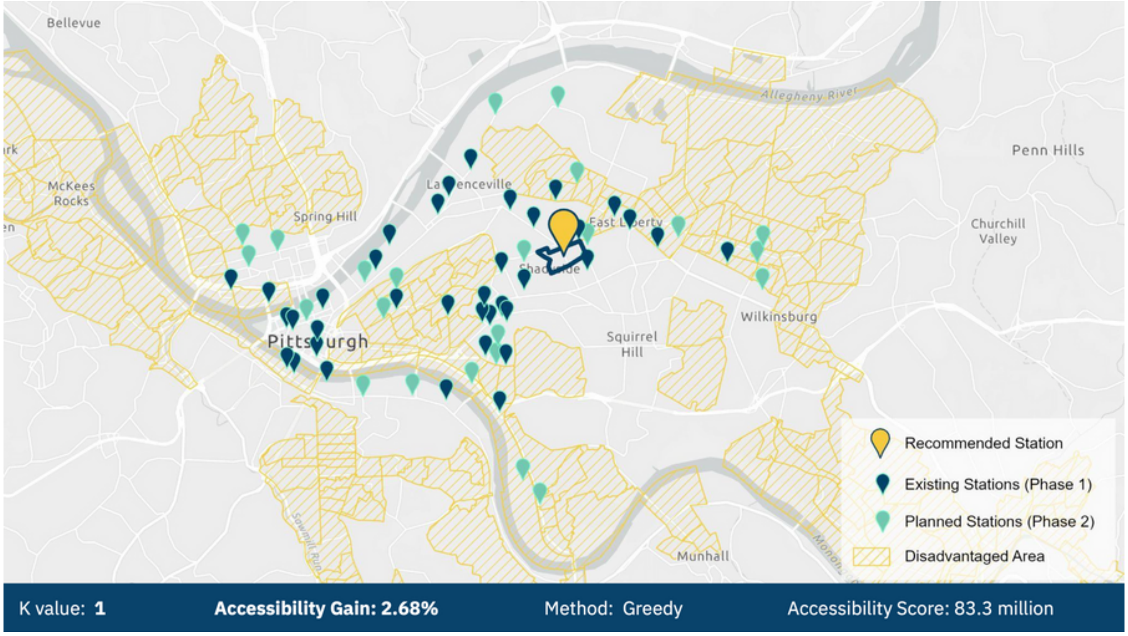
In future runs of the model, POGOH can choose any K value, as long as it is less than the number of candidate stations being considered. In this case we considered 238 candidate stations. The figure above also illustrates that accessibility score does increase with the size of the network. This is what we would expect, because each additional node in a network provides new pathways to opportunities available at every other node.

RECOMMENDATIONS:
OPTIMAL STATION LOCATIONS

Note that our recommendations are for the Census block group level, meaning that the optimal station could fall anywhere within the recommended block group. Because we chose to represent our recommended stations using the center of each block group, the recommended points on the following maps are approximate.

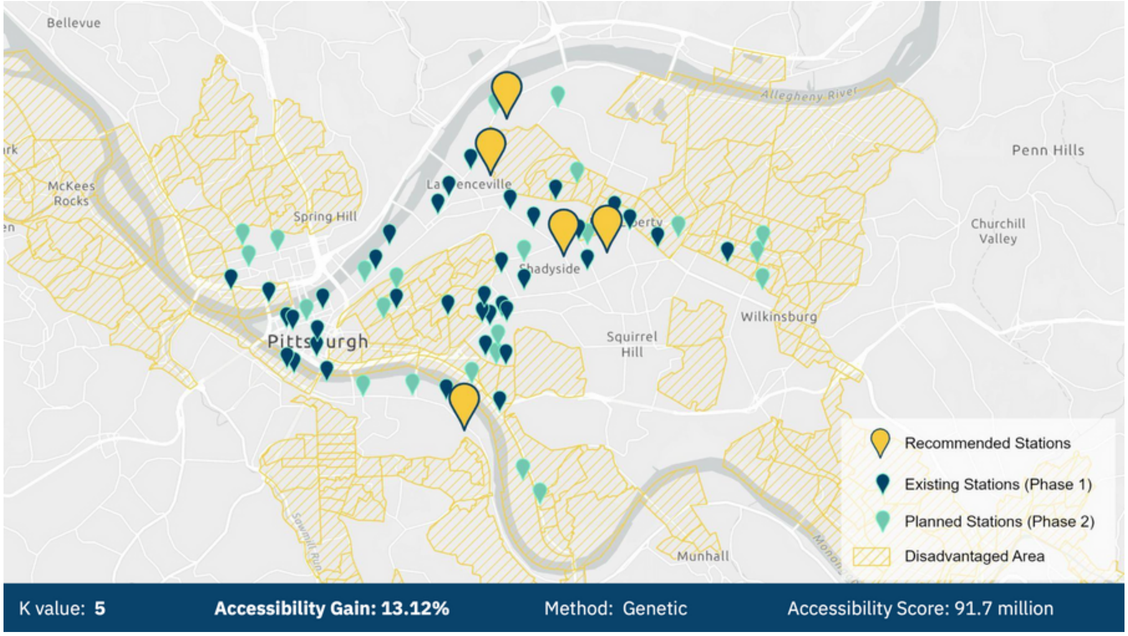
K = 1

The map to the right shows our number one recommended station located between Shadyside and East Liberty. This number one optimal station, which was determined through an exhaustive search, achieves a 2.7% gain in accessibility, compared to the existing network.



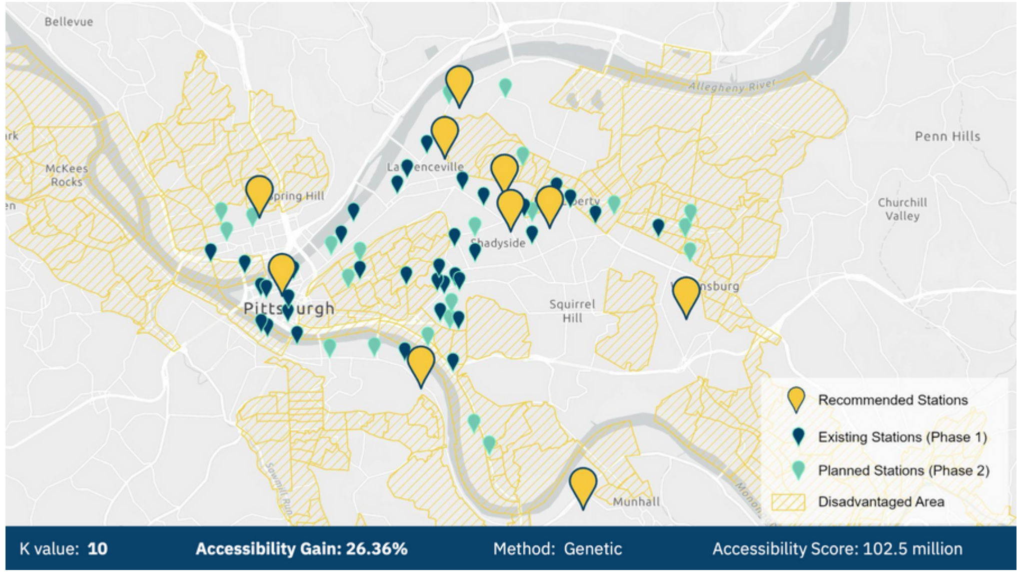
K = 5

When we increase the K value to 5, we see the accessibility score increase to approximately 13% greater than the accessibility score of the existing network. These recommended station locations are produced by the genetic algorithm, which allows us to analyze the network as a whole. The map to the right identifies these optimal stations, which expand further out north of Lawrenceville and onto the South Side.



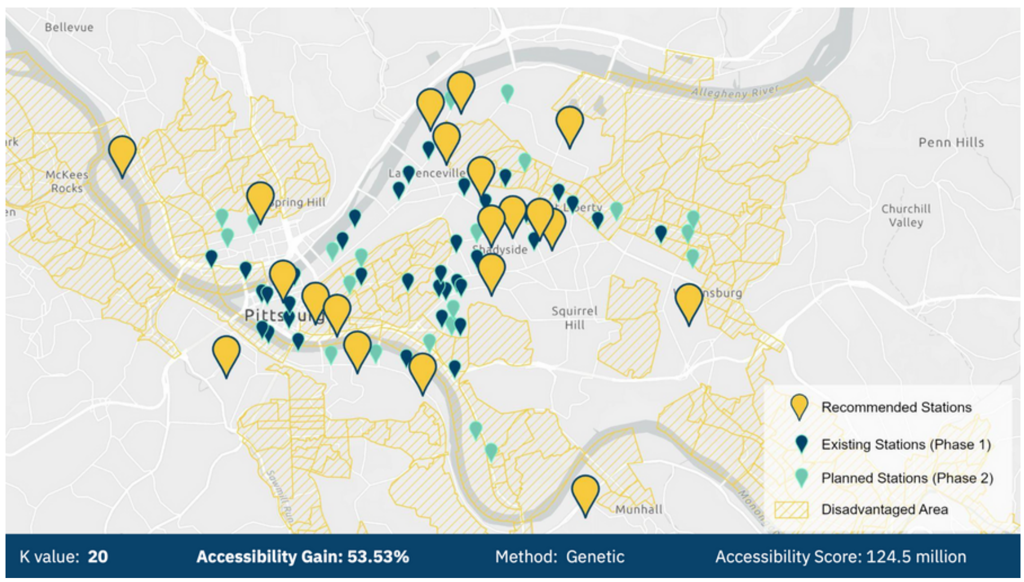
K = 10

Increasing to a K value of 10, the optimal stations expand out geographically, with one farther east near Wilkinsburg, one farther south in Homestead, one on the Northside, and another downtown. Note that the accessibility gain has increased to 26.4%.



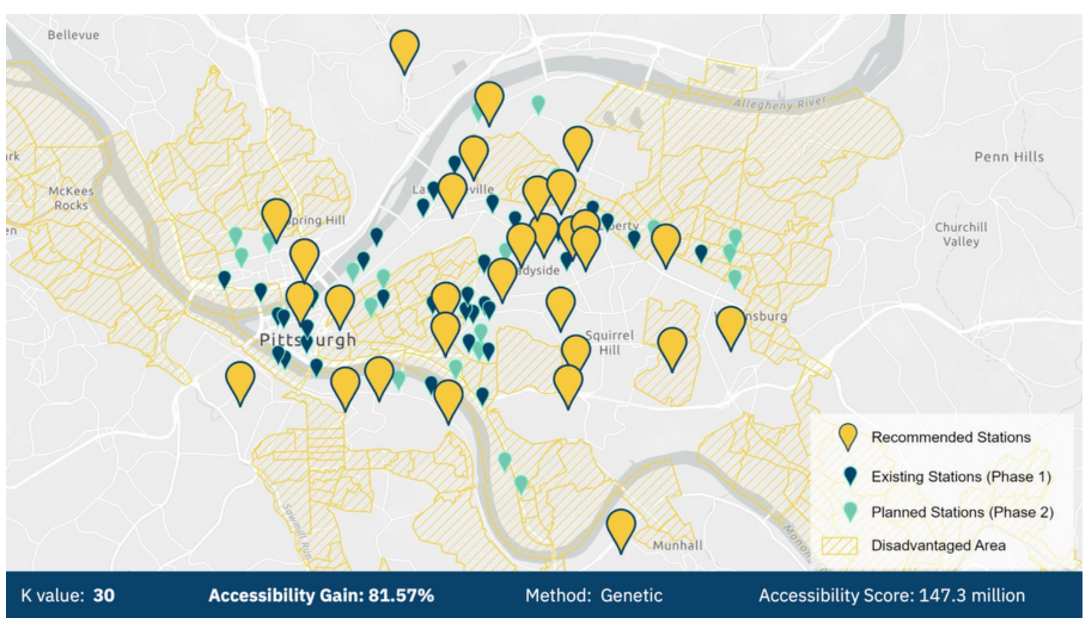
K = 20

To the right we can see that adding 20 stations to the network can increase accessibility by more than 50%. These stations add to the density around downtown and East Liberty, as well as expand into Mt. Washington and McKees Rocks.



K = 30

The map in the bottom right illustrates that adding 30 stations to the network could increase accessibility by more than 80%. Note that these 30 stations do not include every station from the previous map of 20 stations. Because the genetic algorithm optimizes over the entire set of 30 additional stations at once, it is able to find a combination of stations that produces a better solution than just adding on to the previous optimal solution.



These station recommendations are the results of our data-driven model for maximizing accessibility at the block group level, when considering the 238 candidate stations within a 1-mile buffer of existing stations, and for the five different K values. We recommend that Bike Share Pittsburgh integrate this model into their station siting methodology and customize the input parameters. Given different inputs than we used here, the model will likely produce different optimal station configurations.

Takeaway:

We recommend that Bike Share Pittsburgh run this model, discuss with its stakeholders, and implement optimal station recommendations to maximize accessibility, and therefore equity.

RECOMMENDATIONS: ADDITIONAL RECOMMENDATIONS

In addition to our station location recommendations, we have several additional recommendations to propose. First, we recommend that Bike Share Pittsburgh move forward with a collaboration with the Allegheny County Department of Human Services. This partnership can provide rich information about rider demographics, which can in turn be used to better understand past ride data, and better predict future rides.

Next, we recommend that Bike Share Pittsburgh engage with Mobility Justice Members who we identified as having significantly high numbers of trips. Conversations with MJM riders who ride frequently will be able to provide valuable insights regarding network improvements that will benefit disadvantaged riders.

Lastly, we recommend that this data-driven approach be used as just one component of siting new bikeshare stations. The optimal stations produced by the model can be used as a starting point for engaging in conversation with communities.

RECOMMENDATIONS: LIMITATIONS

We did face some challenges with our model. One of the main limitations of our model is the predictive power of the regression, given the limited amount of data available to us at the time of analysis. As the network expands and time goes on, we expect that a greater amount of data will increase the predictive power of the model.

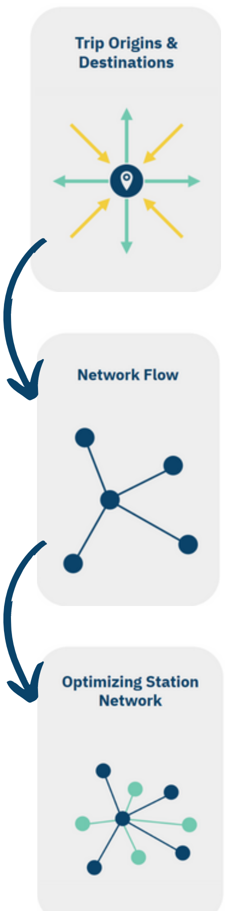
Our model is also limited by the assumption that rider demographics reflect the demographics of the block groups where their trip originated. While this assumption is not uncommon and aligns with current research practice, it does place limitations on the accuracy of our predictions.

FUTURE WORK

Improve Data

Finally, we would like to offer some ideas about future iterations of this work. As previously discussed, re-running the model with more data and with richer data could produce more reliable results. The quantity of data will increase over time and with more stations, and a partnership with Allegheny County could also enrich the data.

We also found that it could be useful to know which rides are taken by University of Pittsburgh students, given that they have access to POGOH bikes through the university. There is also potential to increase the geographic granularity of analysis by zooming in to a level smaller than Census block groups.



Improve Methods

Adjustments can also be made to the methods used in the model. Bike Share Pittsburgh might want to change the list of candidate stations — for example, by excluding block groups in certain areas or targeting block groups in other areas.

Additional objectives can also be incorporated into the optimization. For example, the model could be adjusted to also maximize revenue when generating optimal stations.

Additional constraints can also be added to the optimization. For example, Bike Share Pittsburgh may want to ensure that a certain share of new stations added to the network are in disadvantaged areas. By fine-tuning our tool, the organization could test what the overall network accessibility would be given different siting strategies.

REFERENCES

Anzilotti, E. (2019, June 18). Madison is the first city to go 100% electric with its bike share. Fast Company. https://www.fastcompany.com/90365265/madison-is-the-first-city-to-go-100-electric-for-its-bike-share

Bike PGH. (2022). Bicycling and Walking in Pittsburgh 2022: A snapshot of trends and statistics. BikePGH.org. https://bikepgh.org/wp-content/uploads/2022/02/Cycling-Trends-Snapshot-in-Pittsburgh-2022.pdf

City of Pittsburgh (2018). Pittsburgh Equity Indicators. PittsburghPA.gov. https://pittsburghpa.gov/equityindicators/documents/PGH_Equity_Indicators_2018.pdf

COGO. (n.d.) About CoGo Bike Share. https://cogobikeshare.com/. https://cogobikeshare.com/about

Divvy. (n.d.) Home. Divvy Bikes. https://divvybikes.com/

Divvy. (n.d.) What is a station? A dock? A kiosk? Divvy Help. https://help.divvybikes.com/hc/en-us/articles/360033124372-What-is-a-station-A-dock-A-kiosk

Estimated Average Number of Vehicles per Household, 2017-2021. PolicyMap.

Estimated percent of housing units for which no vehicles are available, 2017-2021. PolicyMap. https://cmu-policymap-com.cmu.idm.oclc.org/newmaps#/

Estimated percent of workers who commuted to work using public transit, 2017-2021. PolicyMap. https://cmu-policymap-com.cmu.idm.oclc.org/newmaps#/

Fan, Zhufeng, and Corey Harper (2022). "Congestion and environmental impacts of short car trip replacement with micromobility modes." Transportation Research Part D: Transport and Environment, Vol., 103. https://doi.org/10.1016/j.trd.2022.103173

Green Cities (2019).Vehicle Ownership in U.S. Cities Data. vanderbilt.edu. https://my.vanderbilt.edu/greencities/files/2019/10/Vehicle-Ownership-in-U.S-2016.pdf

Greenfield, J. (2021, Sep. 24). The reward for expansion and electrification of Divvy is record-breaking ridership. StreetsBlogChicago. https://chi.streetsblog.org/2021/09/24/the-reward-for-expansion-and-electrification-of-divvy-is-record-breaking-ridership/

Herbert, K. (2023, Feb. 24). The 2023 Indego Bike Share Equity Plan is Here. Better Bike Share Partnership. https://betterbikeshare.org/2023/02/24/the-2023-indego-bike-share-equity-plan-is-here/

IndeGo. (n.d.). Media. RideIndego.com. https://www.rideindego.com/about/media/

MadisonBCycle. (n.d.) Community Pass Program. Madison.BCycle.com. https://madison.bcycle.com/nav/in-the-community/community-pass-program

Metrobike. (n.d.). BCycle For All. Austin.BCycle.com. https://austin.bcycle.com/nav-2/bcycle-for-all.

Metrobike. (n.d.). Our Story. Austin.BCycle.com. https://austin.bcycle.com/who-is-metrobike.

Metropolitan Transportation Commission. (n.d.). Bay Wheels Bike Share Program. mtc.ca.gov. https://mtc.ca.gov/operations/traveler-services/bay-wheels-bike-share-program#:~:text=Bay%20Wheels%20brings%20affordable%2C%20accessible,transportation%20around%20the%20Bay%20Area.

Nuebel, K. (2021, July 20). Here's How Pittsburgh is Handling the "Growing Pains" of Spin Scooters. Pittsburgh Magazine. https://www.pittsburghmagazine.com/heres-how-pittsburgh-is-handling-the-growing-pains-of-spin-scooters/

Qian, Xiaodong, Miguel Jaller, and Debbie Niemeier. "Enhancing Equitable Service Level: Which Can Address Better, Dockless or Dock-Based Bikeshare Systems?" Journal of Transport Geography 86 (June 1, 2020): 102784. https://doi.org/10.1016/j.jtrangeo.2020.102784.

Qian, Xiaodong, and Miguel Jaller. "Bikeshare Destination Choices and Accessibility among Disadvantaged Communities." Transportation Research Part D: Transport and Environment 91 (February 1, 2021): 102686. https://doi.org/10.1016/j.trd.2020.102686.

Qian, Xiaodong, and Deb Niemeier. "High Impact Prioritization of Bikeshare Program Investment to Improve Disadvantaged Communities' Access to Jobs and Essential Services." Journal of Transport Geography 76 (April 1, 2019): 52–70. https://doi.org/10.1016/j.jtrangeo.2019.02.008.

Pittsburgh Region (n.d.) Public Transit in Pittsburgh. PittsburghRegion.org. https://pittsburghregion.org/living-in-pittsburgh/transportation/

Pittsburgh Regional Transit (2023). ACCESS Paratransit. RidePRT.org. https://www.rideprt.org/inside-Pittsburgh-Regional-Transit/rider-info/Accessible-Services-and-Features/access-paratransit/

POGOH. (n.d.) Media Kit. POGOH.com. https://pogoh.com/media-kit/

Porter, J. (2019, Jul. 29). Far South Siders Weigh in on the New Electric Divvy Bikes at We Keep You Rollin's Ride. StreetsBlogChicago. https://chi.streetsblog.org/2019/07/29/south-siders-weigh-in-on-the-new-electric-divvy-bikes-at-we-keep-you-rollins-ride/

Redfin. (2023, Jan.). 10 most transit-friendly neighborhoods in Pittsburgh. Redfin.com. https://www.redfin.com/city/15702/PA/Pittsburgh/most-transit-friendly-neighborhoods

ROAM Nebraska. (n.d.) Who We Are. roamshare.org. https://www.roamshare.org/who-we-are

San Francisco Municipal Transportation Agency (n.d.). Bikeshare. sfmta.com https://www.sfmta.com/getting-around/bike/bike-share

Spin. (2021, Aug. 23). How We Plan to Help Make Pittsburgh an Inclusive 15-Minute City. Spin Scooters. https://www.spin.app/blog-posts/how-we-plan-to-help-make-pittsburgh-an-inclusive-15-minute-city

Trouble- Free. (2021, Aug. 3) #53 Genetic Algorithm - Introduction |ML|. https://www.youtube.com/watch?v=bBxUaHJTNNI

Warren, Brent. (2023, Mar. 6). 2022 Was a Record Year for CoGo; Scooter Ridership Still High. Columbus Underground. https://columbusunderground.com/2022-was-a-record-year-for-cogo-scooter-ridership-still-high-bw1/

APPENDIX

TRANSIT EQUITY IN PITTSBURGH:

A data-driven approach

2023

APPENDIX 1:
CURRENT STATE OF TRANSIT

The City of Pittsburgh Department of Mobility and Infrastructure (DOMI) is responsible for transportation services for both people and goods, along with the management of public rights-of-way (which include streets, sidewalks, curbs, and bridges). DOMI is in charge for setting policies and issuing permits for micro-mobility and shared-mobility operators. The mission of the Department is to provide the physical infrastructure that enables residents and visitors to have the social and economic mobility that they deserve.

The Pittsburgh Mobility Collective (PMC), a partnership of transportation service providers, was set up by DOMI to ensure that shared mobility services are deployed equitably, efficiently, and cohesively across Pittsburgh. The PMC leverages public-private partnerships to build a better transportation system that integrates Pittsburgh Regional Transit services, shared electric scooters and bicycles, expanded carshare, and carpool services, using a simple trip planning mobile app and on-street locations known as “Mobility Hubs”.

Pittsburgh Regional Transit (PRT), formerly known as the Port Authority of Allegheny County, is the largest transit agency in southwestern Pennsylvania. It provides public transportation and Access services throughout the City of Pittsburgh and Allegheny County. The Authority operates a total of 102 routes, including buses, light rails, and inclines and serves almost 200,000 riders daily.

Discussed in detail on the following 2 pages, Pittsburghers travel by:

Light rail

Walking

Bus

Bike

Spin Scooter

APPENDIX 2:
CURRENT STATE OF TRANSIT

Light Rail

Pittsburgh has a light rail system, known as the T, which operates two main lines: the Red Line and the Blue Line. The light rail system serves downtown Pittsburgh, the South Hills, and the North Shore. The Pittsburgh light rail system features 27 total stations that span 26.2 miles.

Walking

Pittsburgh is also a great city for getting around on foot. Earning a 63 Average on Walkscore, Pittsburgh is among the most walkable cities in the US, particularly in neighborhoods such as Downtown, Oakland, and Shadyside, where many destinations are within a short walking distance. Over 16,000 Pittsburghers walk to and from their job each day and according to the Make My Trip Count survey 28% of Pittsburgh residents include walking among their top three commute modes.

Buses

PRT operates a bus system that covers a wide area of the county, including Pittsburgh's urban core, as well as outlying suburbs and municipalities. The system operates over 100 routes, including local and express routes, as well as specialty services such as the Free Fare Zone downtown and the Airport Flyer. Almost 84% of all bus routes travel to Downtown Pittsburgh. PRT bus services can be accessed using the handy ConnectCard for faster payments and the fare is $2.75 for unlimited rides within a three hour period.

APPENDIX 1:
CURRENT STATE OF TRANSIT

Bikes

Historically, the number of Pittsburgh residents who use bicycles have not been very high especially given the hilly terrain and less than ideal weather. However, Pittsburgh is now nationally recognized with having the second fastest 20-year growth in bike commuters in the United States. And of the 60 largest US cities, Pittsburgh has the 13th highest number of bike commuters with a total of 14.9% residents using bikes in 2022. People earning less than $30,000 per year accounted for 28% of bike trips. Along with that, 40% of those who don't cycle mentioned that more bike lanes would encourage them to do so. The City of Pittsburgh's Department of Mobility and Infrastructure has designated a series of streets across the city as the official bicycle network.

Spin Scooters

Spin Scooters were launched in July 2021 by Move. Anyone 18 years and older can use Spin Scooters by downloading the Transit App for free. The starting fee for rides is $1 plus 39 cents per minute. In October 2021, Spin and DOMI worked to introduce "Access Zones" throughout the city. These zones, which were identified by DOMI using an equity score, help to identify populations with outsized barriers to accessing transportation. Riders who start their trip within an Access Zone receive a 25% discount off of their total trip fee. The company also has a "Spin Access" program. Pittsburghers who qualify for governmental low-income assistance programs are able to receive a 75% discount on every trip after signing up for Spin Access. Till June 2022 186 Pittsburghers had signed up for Spin Access and taken over 4,000 trips.

Spin users in Pittsburgh ride an average of 1,614 trips per day. The majority of trips are kept short, with 65% lasting less than 10 minutes, and 80% lasting less than 15. In a survey with 2212 Spin users, 35% responded that their scooter trips replaced private vehicle trips - taking approximately 257,000 vehicle miles off the road.

APPENDIX 2:
PHASE I TECHNICAL DETAILS

Phase I estimates the number of trip origins and destinations on a block group level, using two negative binomial regression models. Negative Binomial regressions are generally used when the response variable - in this case origins and destinations - are in the form of counts (i.e., counts of trips originating/ending at a block group). The models provide coefficients for the eight predictor variables that are fed into the model, as well as measures of statistical significance. Refer to the tables to the right.

In order to calculate the predicted trip origins and destinations for a block group, we applied the equation below.

Trip Origin Negative Binomiall Regression Coefficients

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	7.107	1.308	5.432	0.000	4.543	9.671
Total_In_labor_force	0.001	0.001	0.601	0.548	-0.001	0.002
Employment_Rate	-0.008	0.009	-0.868	0.385	-0.026	0.010
bike_path_density	2.623	10.552	0.249	0.804	-18.059	23.304
park_density	-0.015	0.172	-0.085	0.932	-0.352	0.323
num_stations	0.773	0.540	1.430	0.153	-0.286	1.831
pct_young_pop	0.373	1.315	0.284	0.776	-2.204	2.951
num_bus_stops	-0.015	0.036	-0.419	0.675	-0.086	0.056
Disadvantaged	-0.347	0.688	-0.505	0.614	-1.695	1.001

Trip Destination Negative Binomiall Regression Coefficients

	coef	std err	z	P>|z|	[0.025	0.975]
Intercept	7.239	1.308	5.533	0.000	4.675	9.804
Total_In_labor_force	0.000	0.001	0.306	0.760	-0.001	0.002
Employment_Rate	-0.011	0.009	-1.160	0.246	-0.028	0.007
bike_path_density	0.677	10.552	0.064	0.949	-20.005	21.359
park_density	-0.010	0.172	-0.056	0.955	-0.347	0.328
num_stations	0.889	0.540	1.645	0.100	-0.170	1.947
pct_young_pop	0.763	1.315	0.580	0.562	-1.814	3.341
num_bus_stops	-0.015	0.036	-0.406	0.685	-0.085	0.056
Disadvantaged	-0.617	0.688	-0.897	0.370	-1.964	0.731

$$log(\text{Trip Origins/Destinations}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_p X_p$$

Where beta refers to the coefficient, and X refers to the corresponding value of a given block group

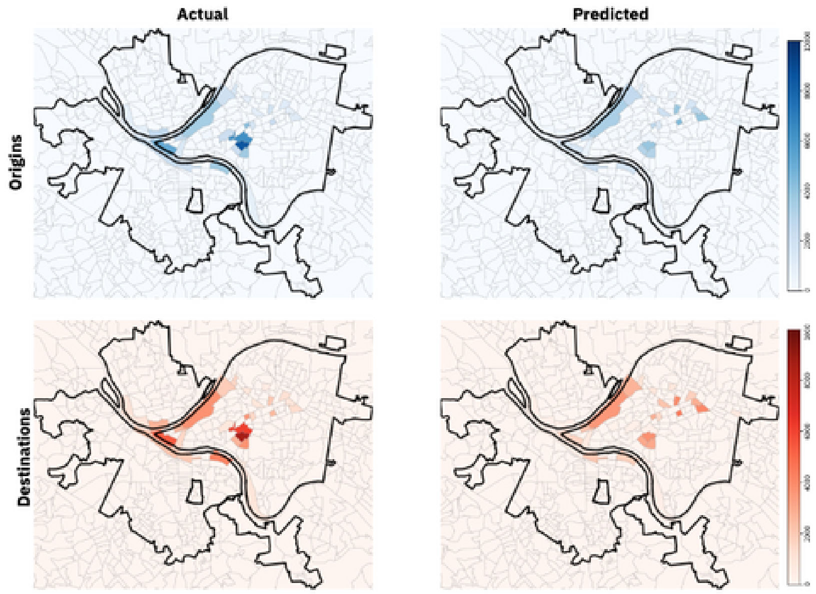
APPENDIX 2:
PHASE I TECHNICAL DETAILS

The results from this phase follow two attempts: 1) Using a scaled Chicago-based model, and 2) using a Pittsburgh-based model.

First, we attempted to predict the trip origins and destinations using a model that was trained by Qian and Jaller using Chicago data. The reasoning behind this is due to the fact that there are only 28 block groups in Pittsburgh that have any POGOH stations, and therefore our number of observations was small, and would not yield any statistical significance. To rectify this, we had tried predicting trip origins and destinations using a model tuned on Chicago data, then scale the predictions by 25% to account for population differences between Chicago and Pittsburgh. Using the scaled Chicago-based model tended to under-predict the number of rides originating and ending in block groups surrounding Oakland and Downtown areas of the city.

	Root Mean Squared Error
Scaled Chicago Model (Origin)	1,972
Scaled Chicago Model (Destination)	1,960
Pittsburgh Model (Origin)	992
Pittsburgh Model (Destination)	1,413

Chicago Model Predictions vs. Actual Trips



Ultimately, we decided that statistical significance was not a necessary feature for our use case, and instead we would focus on predictive power of the prediction models. Using a model trained on Pittsburgh data yielded much better predictions, and lower root mean squared errors than a scaled Chicago model.

APPENDIX 3:
PHASE II TECHNICAL DETAILS

We referred to the research from Qian and Jaller (2021) to estimate spatial distribution of bikeshare rides. This study develops a competing destination model to understand destination choices and behaviors of bikeshare users.

We used the following to calibrate the model:

1. Publicly available bikeshare trip data from POGOH (2022-23)
2. Jobs and points of interests data to evaluate attractiveness of a census block group from the Longitudinal Employer-Household Dynamics (LEHD) dataset
3. Distance and trip times between stations collected through Google API.

We calibrated the model and estimated parameters using the existing data. We then apply these parameters to forecast trip distribution between any two stations in the future based given their attractiveness and distance /time between the stations.

Bikeshare destination choices and accessibility among disadvantaged communities is a case study based on Chicago's bikeshare program, Divvy. This study addresses the lack of understanding of trip destination choices and behaviors for bikeshare users from disadvantaged communities.

The research offers a methodology to analyze spatial patterns bikeshare trips. Following a competing destination model based on a gravity model, this study distributes the trips originated from station A to other stations in the network. It also assesses the effect of improved accessibility on destination choices for different populations and user types.

The study developed a prediction model that maximizes travel benefits for the system while minimizing travel decay functions. Similar to the gravitational force that increases between higher mass objects and decreases with the distance between them, this model predicts that there would be more trips between stations with higher accessibility improvement and fewer trips between stations that are located far from each other.

APPENDIX 4:
PHASE III TECHNICAL DETAILS

Overview

As discussed in the report, we used the genetic algorithm to optimize POGOH's network for accessibility, defined as the change in opportunities (jobs, grocery stores, and more) that someone experiences by moving from one station to another. As a check against the genetic algorithm, we also tested the greedy algorithm. The greedy algorithm is "greedy" because it chooses one station at a time that contributes the greatest accessibility gain to the network. Importantly, it does not account for network effects or synergies that might occur by siting stations that individually may not contribute the greatest accessibility gain, but together create large gains. However, it does have notable advantages: it converges quickly and is deterministic, meaning that it produces the very same list of optimal stations every time that it is run.

The genetic algorithm is "genetic" in nature because it uses evolutionary concepts: selection, crossover, and mutation. The genetic algorithm accounts for network effects and station synergies by optimizing for the whole network at one time. However, it is complex in nature, which can lead to long run times, and is also stochastic, which means that there is a measure of randomness at play, so the optimal station list could shift slightly if rerun. Ultimately, we leveraged the genetic algorithm as our optimization technique because of the importance of network effects.

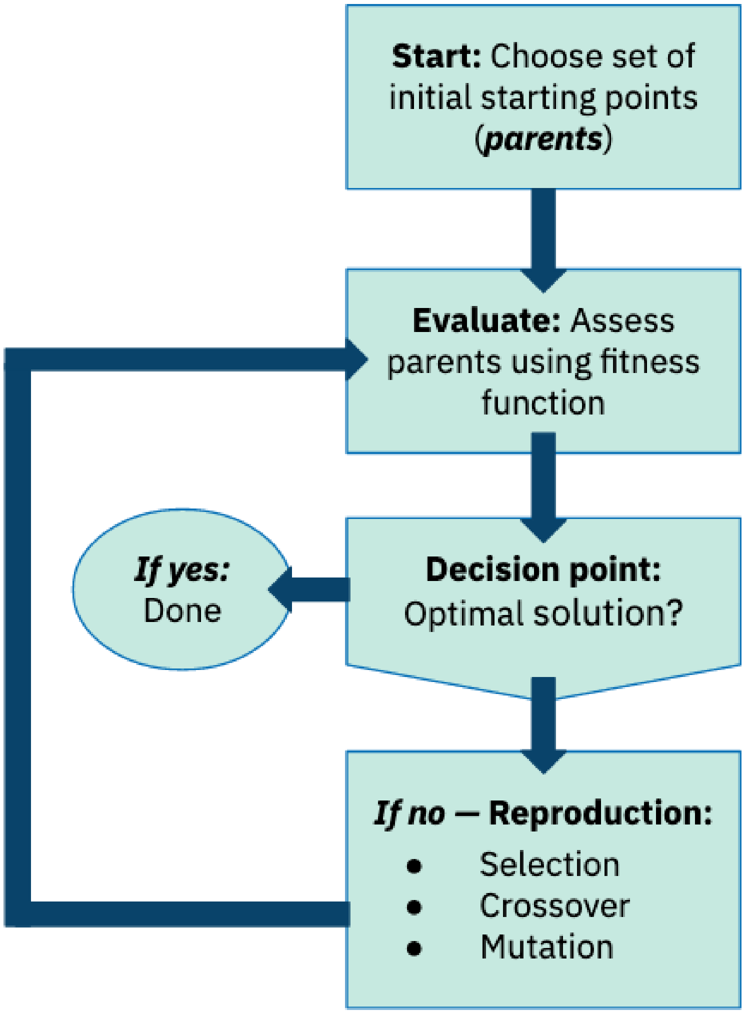
The Genetic Algorithm

The genetic algorithm is the optimization algorithm that we selected, based on Qian, et al (2022), to maximize network accessibility. This algorithm works well for complex problems with large search spaces. It may not select the overall optimal solution, but as we continue to increase the number of runs, the solution value converges toward the global maximum. The overall network accessibility is the sum of the change in accessibility from station i to station j multiplied by the number of trips predicted to take that route (from station i to station j).

$$Maximize: \sum_{i=1}^{238} \sum_{j=1}^{238} Accessibility\ change_{\ i,j} * Number\ of\ trips_{ij}$$

APPENDIX 4:
PHASE III TECHNICAL DETAILS

As mentioned above, the genetic algorithm leverages so-called genetic techniques to converge. These include **selection**, where we select the parent solutions with the highest accessibility gains; **crossover**, where we merge the selected stations at a given point from parent 1 and parent 2; and **mutation**, where we randomly change or remove stations from the child solutions to introduce novel combinations. The last step, in particular, helps to ensure that we do not get trapped in a local minimum but rather work toward the global maximum with the highest accessibility score. A diagram outlining the conceptual flow of the algorithm is shown to the right.

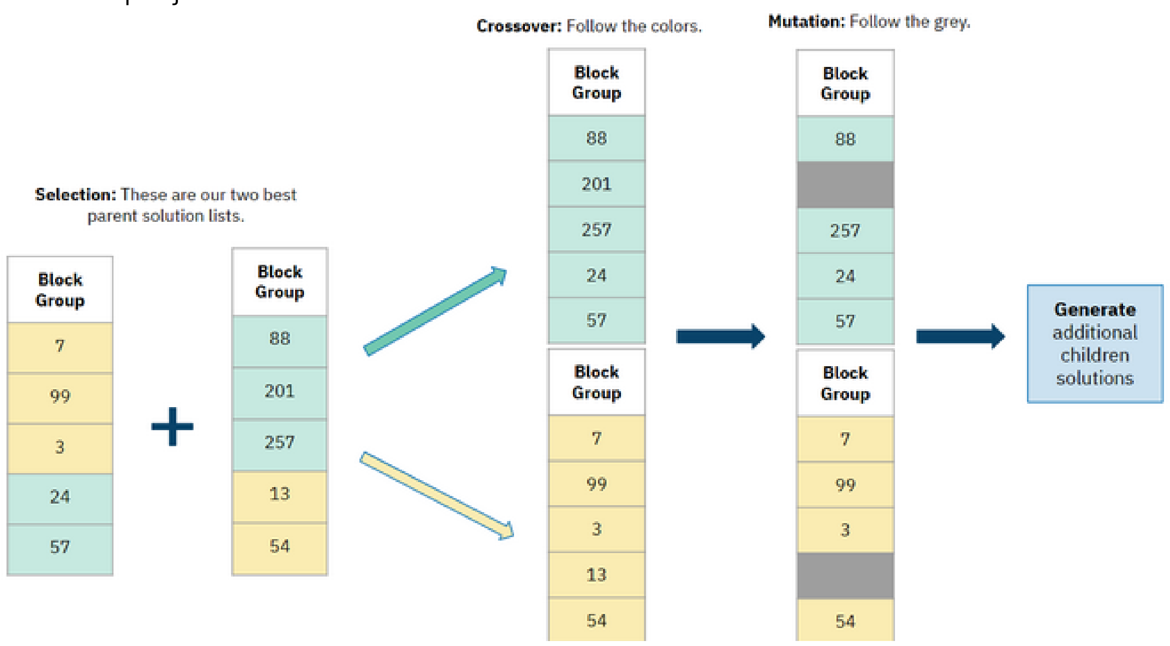


- **Fitness Function:** Maximize Accessibility
- **Genes:** A list of stations that the fitness function has selected as producing the highest accessibility gain.
- **Parent Solutions:** A list of the block groups that the fitness function has selected to host new stations. Parents are selected because they maximize accessibility.
- **Children:** The new solution values that are created by crossing over the lists of block groups from the two parents.
- **Mutations:** Randomly removing stations from the pool.
- **K**: The number of stations to consider adding in a given run. E.g., if K = 5, we would select the 5 best stations to increase network accessibility.

APPENDIX 4:
PHASE III TECHNICAL DETAILS

K is an important parameter to highlight. K is the number of stations that we consider adding in a given run of the algorithm. For example, if Bike Share Pittsburgh wanted to consider what the accessibility gain would be if 5 new stations were added, K would be set to 5. We tested a number of different K values to explore what the increase in network accessibility would be. This may be especially useful for Bike Share Pittsburgh to understand the potential impact of its network expansions.

Below, we have provided a diagram that walks through the different steps of the genetic algorithm. First, we **select** the "fittest" parents with the highest accessibility gains. The parents are simply lists of stations that generate the most accessibility for the network. Next, we break the lists into two to enable **crossover.** We elected to do a one-point crossover, where we split the stations genes at one point, but it is possible to do two-, three-, or even multi-point crossover. In the diagram below, we crossed over the station genes after the third gene in the list. The teal genes (genes 4 and 5 from Parent A and genes 1, 2, and 3 from Parent B) are merged into one child solution, and the yellow genes are merged into another child solution. Lastly, we introduce **mutations** into the child solution station lists. This simply means randomly removing stations from the child station lists. In our algorithm, we set the mutation rate to 10%, but this is a tunable parameter that can be changed in future iterations of the project.



APPENDIX 4:
PHASE III TECHNICAL DETAILS

We continue to generate additional children solutions using this strategy, creating a **generation** of solutions. We then select the fittest solutions with the highest accessibility scores from the child generation. We also keep the fittest parent solutions in the event that more optimal solutions, with higher accessibility scores, were not produced in the new child generation. **The beauty of the genetic algorithm is that as we generate more child solutions and increase the number of child generations to produce**, we converge toward a global maximum. Given the large search space of the problem, this is an incredibly powerful method to leverage. By running the algorithm with a very large number of child generations (e.g., 1,000), POGOH could further increase network accessibility by finding more optimal solutions. (Given our computational resources, we have set the number of child generations at 100.)

Lastly, we compare the results of the genetic algorithm and the greedy algorithm. First, we see that the two algorithms produce very similar results. This buttresses our findings: that both algorithms, which leverage disparate techniques, produce near-identical results implies solution stability. Second, we see that increasing the K parameter — the number of stations by which to increase the network — results in effectively linear increases in accessibility. Again, this is true both of the greedy and genetic algorithm. Because demand is limited, we would expect that this trend would likely taper off, resulting in a horizontal asymptote. However, this is also an area for future work: we can set constraints so as to hem in ridership demand and not assume infinite increases in ridership — and accessibility — given network expansion.